

COMPOSIÇÃO POPULACIONAL POR RAÇA/COR E DISPARIDADE EM TAXAS DE MORTALIDADE POR COVID-19 EM SÃO PAULO: UMA ANÁLISE ESTATÍSTICA E ESPACIAL

Agnes Silva de Araujo

Pontifícia Universidade Católica do Paraná (PUCPR), São Paulo, SP, Brasil

E-mail: agnes.silva.araujo@gmail.com

Carina Sernaglia Gomes

Senac - campus Jabaquara, São Paulo, SP, Brasil

E-mail: ambiente.carina@gmail.com

Resumo

O impacto desigual da pandemia de COVID-19 no Brasil pode ser observado nas maiores taxas de mortalidade entre a população negra. Este artigo apresenta os resultados de uma análise exploratória da relação estatística e espacial entre as taxas de mortalidade por COVID-19 e a composição populacional por raça/cor na cidade de São Paulo. Os resultados apontam que existe uma relação significativa e positiva entre a composição populacional e as taxas de mortalidade por COVID-19 (R^2 de 0,54), sendo os efeitos da maior proporção de negros, mais significativo. Testes de dependência espacial apontam que os modelos de estatística espacial são os mais adequados para os dados em questão. Os resultados do *Spatial Error* e GWR foram 0,78 e 0,84 respectivamente. O que significa que o modelo de estatística local é o que possui o melhor ajuste. Os resultados apontam que os efeitos desiguais do COVID-19 sobre a população negra refletem as desigualdades sociais e territoriais existentes antes da pandemia.

Palavras-chave: COVID-19; composição populacional por raça/cor; taxas de mortalidade; análise estatística; análise espacial.

POPULATION COMPOSITION BY RACE/COLOR AND DISPARITY IN MORTALITY RATES BY COVID-19 IN SÃO PAULO: A STATISTICAL AND SPATIAL ANALYSIS

Abstract

The uneven impact of the COVID-19 pandemic in Brazil can be seen in the higher mortality rates among the black population. This paper presents the results of an exploratory analysis of the statistical and spatial relationship between COVID-19 mortality rates and population composition by race/color in the city of São Paulo. The results indicate that there is a significant and positive relationship between population composition and mortality rates from COVID-19 (R^2 of 0.54), with the effects of a higher proportion of black people being more significant. Spatial dependence tests show that spatial statistics models are the most suitable for the data in question. The Spatial Error and GWR results were 0,78 and 0,84, respectively. Which means that the local statistical model is the one that has the best fit. The results show that the unequal effects of COVID-19 on the black population reflect the social and territorial inequalities that existed before the pandemic.

Key words: COVID-19; population composition by race/color; mortality rates; statistical analysis; spatial analysis.

COMPOSICIÓN DE LA POBLACIÓN POR RAZA/COLOR Y DISPARIDAD EN LAS TASAS DE MORTALIDAD POR COVID-19 EN SÃO PAULO: UN ANÁLISIS ESTADÍSTICO Y ESPACIAL

Resumen

El impacto desigual de la pandemia de COVID-19 en Brasil se puede ver en las mayores tasas de mortalidad entre la población negra. Este artículo presenta los resultados de un análisis exploratorio de la relación estadística y espacial entre las tasas de mortalidad por COVID-19 y la composición de la población por raza/color en la ciudad de São Paulo. Los resultados indican que existe una relación significativa y positiva entre la composición de la población y las tasas de mortalidad por COVID-19 (R^2 de 0,54), siendo más significativos los efectos de una mayor proporción de personas de raza negra. Las pruebas de dependencia espacial muestran que los modelos estadísticos espaciales son los más adecuados para los datos en cuestión. Los resultados de error espacial y GWR fueron 0,78 y 0,84 respectivamente. Lo que significa que el modelo estadístico local es el que tiene mejor ajuste. Los resultados muestran que los efectos desiguales del COVID-19 sobre la población negra reflejan las desigualdades sociales y territoriales que existían antes de la pandemia.

Palabras-clave: COVID-19; composición de la población por raza/color; tasas de mortalidad; análisis estadístico; análisis espacial.

Introdução

Inúmeras pesquisas têm relatado e discutido os impactos desiguais da pandemia de COVID-19 nos diferentes grupos populacionais. As taxas de infecções, hospitalizações e mortes variaram, levantando questões sobre quais grupos populacionais enfrentam maiores riscos e por quê. Evidências apontam que pessoas negras estão contraindo SARS-CoV-2 em taxas mais altas e têm maior probabilidade de morrer (MAROKO; NASH; PAVILONIS, 2020; YANCY, 2020). Silva, Morais e Santos (2020) demonstram que apesar da população branca ser a que possui maior números de internação, a população negra possui a maior taxa de mortalidade no Brasil.

O distanciamento social, a restrição de circulação de pessoas, uso de máscaras e medidas de higiene pessoal são estratégias inviáveis para milhões de pessoas que vivem em situações habitacionais precárias e em territórios caracterizados pela maior vulnerabilidade social. Estudos internacionais apontam que a vulnerabilidade social é um fator determinante para a incidência do vírus e maiores taxas de mortalidade (KARAYE; HORNEY, 2020, MAH; ANDREW, 2022; MACKKEY et al., 2021, WANG et al., 2022). A relação entre maior vulnerabilidade social, maior proporção da população negra e mortes em Chicago foi analisada por Kim e Bostwick (2020). Os autores apontam que os efeitos desproporcionais do COVID-19 nas comunidades afro-americanas são um reflexo da desigualdade racial e da exclusão social que existiam antes da pandemia de COVID-19 (KIM; BOSTWICK, 2020).

O Brasil é um país sabidamente desigual e no qual prevalece um racismo estrutural, fato esse que se torna um agravante, porque a população que se autodeclara negra é, frequentemente, a população mais social e economicamente vulnerável (SANTOS et al. 2020) e segregada (FRANÇA, 2016). A vulnerabilidade social vincula-se à noção de

precariedade nas condições de vida da população, que envolve o âmbito do trabalho, da educação, da renda, do saneamento, da saúde, da habitação e da mobilidade urbana (KOWARICK, 2002).

Uma relação espacial e estatística positiva entre a incidência de COVID-19 e o IDH dos municípios do Estado do Ceará, demonstram que os municípios com os menores índices foram os que apresentaram as maiores taxas de incidência de COVID-19 (MACIEL, CASTRO-SILVA, FARIAS, 2020). Em Fortaleza, observou-se uma relação entre as áreas mais socialmente vulneráveis e a incidência de COVID-19 (CESTARI et al., 2021). Baqui et al. (2020), observou uma relação entre maior mortalidade por COVID-19 e as regiões no Brasil com menores índices desenvolvimento socioeconômico e maior ocorrência de comorbidades. Os resultados das pesquisas de Raymundo et al. (2021) corroboram com os de Baqui et al. (2020). Andrade et al. (2020), apontam para uma relação entre a incidência de COVID-19 e áreas de maior densidade populacional e menor nível socioeconômico. No Rio de Janeiro, foi observado uma maior incidência de COVID-19 em regiões de alta renda, o que sugere um acesso desigual à testagem, com a população mais rica tendo maior acesso (RAFAEL et al., 2020).

Conhecimento acerca dos determinantes sociais e espaciais é essencial para identificar áreas com maior potencial de infecções resultantes em mortes, áreas estas que devem ser priorizadas em termos de alocação de recursos humanos e financeiros, além de preparação do sistema de saúde para tratamento de casos graves. Ademais, podem apontar para possíveis políticas públicas de cunho social que podem contribuir para minimizar os efeitos desiguais de mortalidade por COVID-19.

Nessa perspectiva, este artigo visa conduzir uma análise exploratória da relação estatística e espacial entre as taxas de mortalidade por COVID-19 e a composição populacional por raça/cor em São Paulo. A hipótese norteadora desta é a de que existe uma correlação espacial positiva entre maior porcentagem de população negra e a taxa de mortalidade por COVID-19 por distrito em São Paulo.

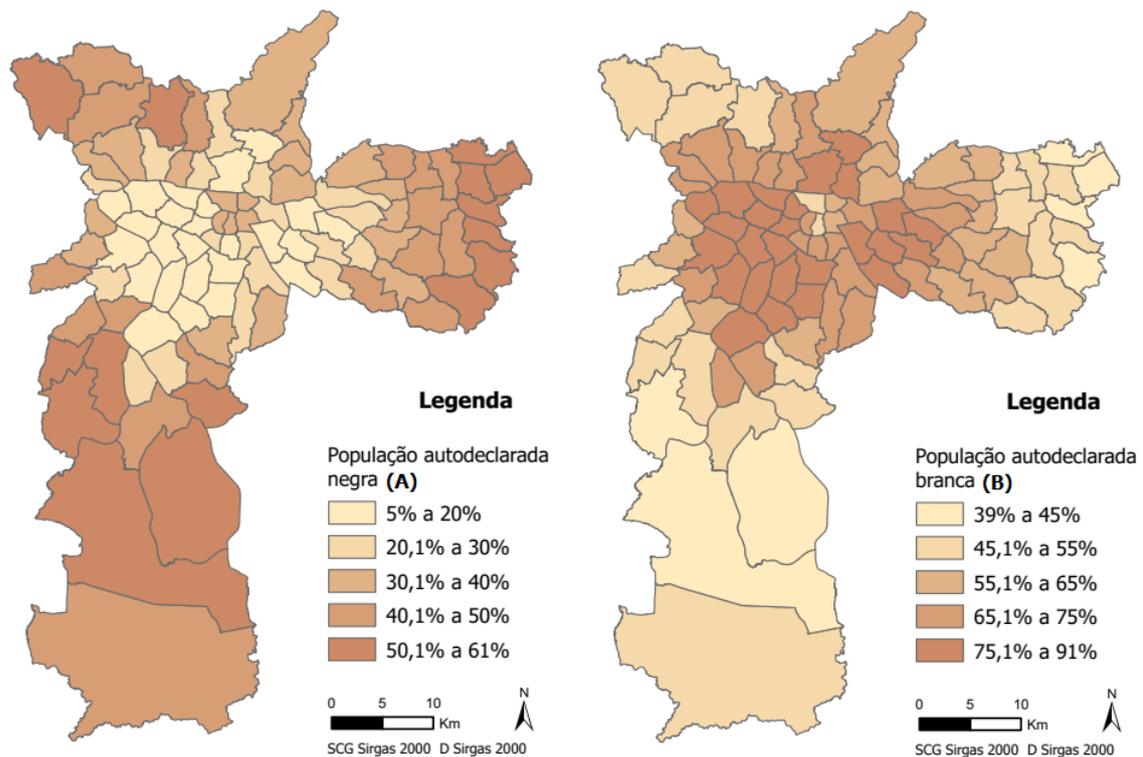
Material e Métodos

Área de estudo

A cidade de São Paulo possui cerca de 11.250.000 habitantes (IBGE, 2010). Desses, aproximadamente 37% pertenciam à população autodeclarada negra (pretos e pardos)

(IBGE, 2010). Como pode ser observado na figura 01, existe um padrão de centro-periferia na concentração da população negra em São Paulo, com as maiores proporções nas áreas periféricas. O oposto pode ser dito em relação à população branca.

Figura 01. Porcentagem da população autodeclarada negra (A) e branca (B) por distrito em São Paulo.



Fonte: IBGE (2010).

No final da década de 1970, pesquisas que relacionam processos sociais aos macro padrões espaciais observados em São Paulo também os descreviam como centro-periferia, onde as elites ocupavam as localizações próximas ao centro, equipadas de serviços e infraestruturas urbanas, enquanto a população com nível socioeconômico mais baixo e mais socialmente vulnerável se localizava nas periferias distantes e desprovidas de infraestruturas urbanas (BONDUKI; ROLNIK, 1982; CAMARGO, 1976; KOWARICK, 1979; MARICATO, 1979; SANTOS, 1978).

Posteriormente, surgiram pesquisas que buscavam explicar o processo de sobreposição, sob uma perspectiva micro escalar, dos novos padrões espaciais aos já consolidados em períodos anteriores (macro padrão centro-periferia), apontando para um processo de aproximação geográfica entre os diferentes grupos socioeconômicos, mantendo

a separação social, através de equipamentos de segurança privada e habitações em estilo condominial (CALDEIRA, 2000; COY; PÖHLER, 2002).

Mais recentemente pesquisas apontam que, com a diversificação das habitações em estilo condominial, os espaços dos grupos de nível socioeconômico mais elevado se tornaram mais homogêneos, em decorrência da multiplicação de condomínios de alto padrão nestas mesmas localidades e, a periferia se tornou mais heterogênea e fragmentada, com a presença de outros grupos socioeconômicos além dos de nível socioeconômico mais baixo (FEITOSA et al., 2021; MARQUES, 2016). No entanto, de forma geral, no que tange à macro escala, evidências apontam que os padrões espaciais são resilientes e há uma manutenção dos padrões observados em períodos anteriores no que concerne a metrópole paulista (FEITOSA et al., 2021; MARQUES, 2016).

Apesar das pesquisas sobre os padrões espaciais da metrópole se referirem a grupos socioeconômicos, é plausível associar os territórios dos grupos de nível socioeconômico mais baixo com a maior proporção da população negra (figura 1). Ademais, atualmente, a proporção de negros em situação de pobreza e extrema pobreza é o dobro da de brancos, a proporção das pessoas negras desocupadas e subocupadas é de cerca de dois terços do total, uma maior proporção da população negra encontra-se em trabalhos informais e ganham salários mais baixos em comparação com os brancos (IBGE, 2019). A desigualdade também aparece nos indicadores de moradia com quase o triplo de pessoas negras vivendo em condições precárias de moradia em comparação com a população branca (IBGE, 2019).

França (2016) em análise sobre segregação por classe social e raça/cor na metrópole paulista, conclui que mesmo em categorias ocupacionais semelhantes, os negros estão mais segregados que os brancos na metrópole paulista e chega a conclusões similares as de Telles (2012) concluindo que o padrão de segregação residencial por raça/cor e classe social parece estar fortemente relacionado com a estrutura de estratificação social no Brasil.

Dados

Para testar a hipótese norteadora desta pesquisa, utilizou-se três variáveis agregadas por distrito. É importante mencionar que se trata de um estudo ecológico exploratório e analítico. As variáveis e as respectivas fontes são apresentadas no Quadro 1, abaixo.

Quadro 01. variáveis utilizadas e suas respectivas fontes.

Variáveis utilizadas	Fonte
Taxa padronizada de mortalidade por COVID-19	Secretaria Municipal de Saúde (SMS) do município de São Paulo – Painel COVID-19.
Porcentagem da população autodeclarada negra	IBGE (2010). Censo demográfico.
Porcentagem da população autodeclarada branca	

Fonte: Os autores (2022).

Em relação à taxa de mortalidade, sua padronização é feita a partir do cálculo do número de óbitos por COVID-19 em 100 mil habitantes por distrito e é ajustada por perfil etário (SMS, s/d). Os valores referem-se aos óbitos confirmados por COVID-19 acumulados desde o início do monitoramento até 22 de março de 2022.

A porcentagem da população autodeclarada negra corresponde à soma das porcentagens identificadas para as populações autodeclaradas preta e parda. O mesmo pode ser afirmado a respeito da população autodeclarada branca.

A partir da seleção das variáveis, foram conduzidos testes estatísticos, de dependência espacial e aplicou-se regressões espaciais globais e locais. A próxima seção apresentará detalhes sobre os métodos.

Regressão estatística

De forma geral, uma regressão estatística é uma técnica que busca prever e explicar o valor de uma variável usando os valores de outra ou outras variáveis, por convenção são denominadas variáveis dependente (Y) e a independente ou explicativa (X) (ALLEN, 1997). Uma regressão simples permite que seja estimada a forma dessa relação, quantificar a precisão com que uma variável explica a outra e possibilita afirmar se a relação entre as variáveis é estatisticamente significativa (ALLEN, 1997).

Nesta pesquisa, buscou-se estimar a relação entre as variáveis: taxa padronizada de mortalidade por COVID-19 (Y) e porcentagem da população negra (X). A Eq. (1) que descreve a relação estatística entre as variáveis foi:

$$Y = \beta\theta + \beta_1.X + \varepsilon \quad (1)$$

Onde: Y representa a taxa padronizada de mortalidade por COVID-19 (variável dependente), β são os parâmetros de interceptação e inclinação do modelo de regressão, X representa a porcentagem da população negra (variável independente) e ε é a soma dos resíduos no modelo.

Testes de dependência espacial

Testes de dependência espacial entre as variáveis foram conduzidos para verificar se a associação estatística entre as variáveis possui uma componente espacial. Aplicou-se: Índice de Moran e os Multiplicadores de Lagrange (*Spatial Lag e Error*).

O Índice de Moran computa a autocorrelação espacial ou o nível de interdependência espacial de uma variável a e o quantifica entre 1 e -1 (ANSELIN, 1995). Valores próximos a positivo ou negativo 1 significam que foi constatada a presença de autocorrelação espacial, ou seja, ou a variável está perfeitamente dispersa (-1) ou agrupada (+1) e, se os testes de pseudo significância determinarem significância estatística, é necessário escolher um modelo que agregue uma componente espacial que melhor descreva a relação de dependência entre as variáveis (ANSELIN, 1988). O I de Moran é computado através da Eq (2):

$$I = (N/W) \frac{\sum \sum w_{ij} (X_i - \bar{X})(X_j - \bar{X})}{\sum (X_i - \bar{X})^2} \quad (2)$$

Onde: N é o número de unidades espaciais indexadas por i e j , W é a soma de todas as unidades vizinhas w_{ij} , X são os valores da variável analisada, \bar{X} é a média da variável x e w_{ij} é a matriz de pesos espaciais de vizinhança ou proximidade.

O teste dos Multiplicador de Lagrange compara, de forma hipotética, os modelos espaciais com a regressão convencional, executando regressões nos resíduos das observações, comparando-os com os resíduos das observações na vizinhança (ANSELIN, 1988). O teste de significância dos Multiplicadores de Lagrange aponta qual modelo espacial é mais adequado para as variáveis em questão.

De forma geral, a interpretação dos resultados dos testes aponta para a violação de premissas de que as observações não são independentes espacialmente e para a necessidade

de se aplicar um modelo mais complexo, que inclua explicitamente, a componente espacial das variáveis. Para tanto, aplicam-se modelos globais ou locais de regressão espacial.

Regressões espaciais globais e locais

A diferença entre os modelos de regressão espacial global e local é que os modelos globais incluem um parâmetro para capturar a estrutura espacial na área de estudo como um todo. Enquanto o *Spatial Lag* (SAR - *Spatial Autoregressive Model*) atribui a autocorrelação espacial a variável resposta Y e o *Spatial Error* (CAR - *Conditional Autoregressive Modeling*), atribui a autocorrelação espacial aos resíduos (erro) do modelo.

A premissa do *Spatial Lag* é que a variável Y_i é afetada pelos valores da variável resposta nas áreas vizinhas a i . Pode ser computado pela Eq. (4):

$$Y = (\rho)Wy + X(\beta) + \varepsilon \quad (4)$$

Onde ρ é o coeficiente espacial autoregressivo ou medida de autocorrelação espacial, W é a matriz de proximidade ou vizinhança espacial e Wy expressa a dependência espacial em Y .

A premissa do *Spatial Error* é que os efeitos espaciais são um ruído, atribuindo a autocorrelação espacial ao erro do modelo, podendo ser computado pela Eq. (5 e 6) a seguir:

$$Y = X(\beta) + \varepsilon \quad (5)$$

$$\varepsilon = (\rho)W\varepsilon + \xi \quad (6)$$

Onde $W\varepsilon$ é o erro com efeitos espaciais, ρ é a medida de autocorrelação espacial e ξ é a componente do erro com variância constante e não correlacionada.

A determinação de qual modelo de regressão espacial global utilizar está relacionada aos resultados dos testes de dependência espacial e fortemente associado aos resultados dos testes de significância dos Multiplicador de Lagrange. Porém, ambos partem do pressuposto

de que o processo analisado é estacionário e pode ser capturado através de um único parâmetro. Alguns fenômenos, no entanto, não apresentam estacionariedade, variando de forma abrupta no espaço.

Para lidar com a heterogeneidade espacial, com fenômenos não estacionários e com relações regionalizadas entre as variáveis dependente e independente, Brunson, Fotheringham e Charlton (1996), propuseram a Regressão Geograficamente Ponderada (traduzido do original *Geographically Weighted Regression - GWR*). É considerada uma regressão espacial local porque ajusta um modelo de regressão para cada ponto do conjunto de dados, ponderando as observações em função da distância de um ponto centroidal.

O pressuposto do GWR é que informações mais próximas do ponto de regressão têm maior probabilidade de influenciá-lo (FOTHERINGHAM; BRUNSDON; CHARLTON, 2002). Essa ponderação é feita pela função *kernel* espacial que pode ser fixa ou adaptativa. O *kernel* usa distância entre dois pontos geográficos e um parâmetro da largura da banda para determinar os pesos das variáveis entre essas duas regiões, que é inversamente relacionado à distância geográfica (FOTHERINGHAM; BRUNSDON; CHARLTON, 2002). A Eq. (6) representa um GWR:

$$Y_i = \beta\theta(ui, vi) + \beta_1(ui, vi).X_i + \epsilon_i \quad (6)$$

Onde Y_i representa taxa padronizada de mortes por COVID-19 no distrito i (variável dependente); β são os parâmetros de intercepção e inclinação do modelo de regressão a serem estimados respectivamente nas localizações θ e 1; X_i representa a porcentagem de população negra (variável independente) no distrito i ; ϵ_i é a soma dos resíduos do modelo de regressão local em i .

Para determinar um valor apropriado para o parâmetro de número de vizinhos na função *kernel* utilizou-se uma técnica automatizada de otimização, denominada pesquisa *Golden Section* (BRUNSDON; FOTHERINGHAM; CHARLTON, 1996). Para o conjunto de dados, foi determinado 30 vizinhos como parâmetro otimizado.

Para avaliar qual modelo de regressão mais adequado para modelar a relação entre as variáveis adotou-se o coeficiente de determinação (R^2) e o critério de informação de Akaike (AIC). O modelo mais adequado é o que possui melhor ajuste entre as variáveis, representado pelos valores elevados de R^2 e o mais simples, representado pelos valores mais baixos de

AIC. A comparação entre os valores de R^2 e AIC são tradicionalmente utilizados para comparar modelos de regressão estatística (ALLEN, 1997).

Resultados e Discussões

Os resultados da regressão clássica são apresentados na tabela 01.

Tabela 01. Resultados dos testes de regressão linear simples entre as variáveis independentes estudadas e a taxa padronizada de mortalidade por COVID-19.

Parâmetros analisados	População autodeclarada negra	População autodeclarada branca
R^2	0,524303	0,472711
R^2 ajustado	0,519242	0,467101
AIC	1041.03	1050.92
Probabilidade global (F de significação)	0,000	0,000
Probabilidade	0,000	0,000

Fonte: Os autores (2002).

Os resultados indicaram que existe uma correlação positiva entre as variáveis independentes analisadas e a taxa padronizada de mortalidade por COVID-19. Os resultados apontaram que a variável “população autodeclarada negra” explica cerca de 52% a variância da taxa de mortalidade. Já para população branca esse valor diminui para aproximadamente 47%.

A fim de analisar a influência das duas variáveis independentes de maneira combinada sobre a taxa padronizada de mortalidade por COVID-19, foi realizada uma regressão linear múltipla. Considerando que o somatório das porcentagens das populações autodeclarada negra e autodeclarada branca refletem aproximadamente o total da população dos distritos. Os resultados são apresentados na tabela 02.

Tabela 02. Resultados do teste de regressão linear múltipla entre as variáveis independentes: porcentagem da população autodeclarada negra e branca e a variável dependente: taxa padronizada de mortalidade por COVID-19.

Parâmetros analisados	População autodeclarada negra e população autodeclarada branca
R^2	0,542929
R^2 ajustado	0,53310
AIC	1039.2

Probabilidade global (F de significação)	0,000
Probabilidade variável independente População autodeclarada negra	0,00028
Probabilidade variável independente População autodeclarada branca	0,05458

Fonte: Os autores (2022).

Ao considerar as duas variáveis como explicativas da taxa de mortalidade, observou-se um melhor ajuste de modelo em função de um valor mais elevado do R^2 , assim como um menor valor de AIC. As duas variáveis independentes combinadas explicaram aproximadamente 53% da variância da taxa padronizada de mortalidade por COVID-19. Destaca-se que apesar de ambas as variáveis independentes apresentarem significância estatística para os testes de probabilidade individuais, o valor do *p-value* encontrado para a variável população autodeclarada negra (0,00028) indica que há significância estatística ao se considerar um intervalo de confiança ainda maior, cerca de 99%, o que não ocorre para a variável relacionada à população autodeclarada branca, cujo limite do intervalo de confiança é de 95% para significância estatística. Esses valores podem ser observados na tabela 03,

Tabela 03. Resultados do teste de regressão linear múltipla entre as variáveis independentes: porcentagem da população autodeclarada negra e porcentagem da população autodeclarada branca e a constante: taxa padronizada de mortalidade por COVID-19.

Variável	Coefficiente	Erro padrão	Estatística-t	Probabilidade
CONSTANTE	-193.078	206.212	-0.936309	0.35154
População autodeclarada negra	7.61372	2.01429	3.77986	0.00028
População autodeclarada branca	4.26007	2.18826	1.94679	0.05458

Fonte: Os autores (2022).

Ainda, a partir da tabela 03 é possível identificar os valores dos coeficientes que indicam como a variável dependente varia ao se estabelecer uma das variáveis independentes como controladas. Ao considerar a variável população autodeclarada negra como controlada, a cada aumento de unidade na variável população branca, a taxa padronizada de mortalidade por COVID-19 aumenta em 4,26. Da mesma maneira, ao considerar a variável população autodeclarada branca como controlada, cada aumento de unidade da variável população

autodeclarada negra faz com que a taxa da variável dependente aumente 7,61, indicando maior influência desta variável independente no modelo.

A fim de se obter um modelo com melhor ajuste, foram realizados testes de dependência espacial. Os testes indicaram a existência de autocorrelação de todas as variáveis abordadas, apontando para a importância da consideração do espaço nas análises de regressão. Os resultados são apresentados na tabela 04.

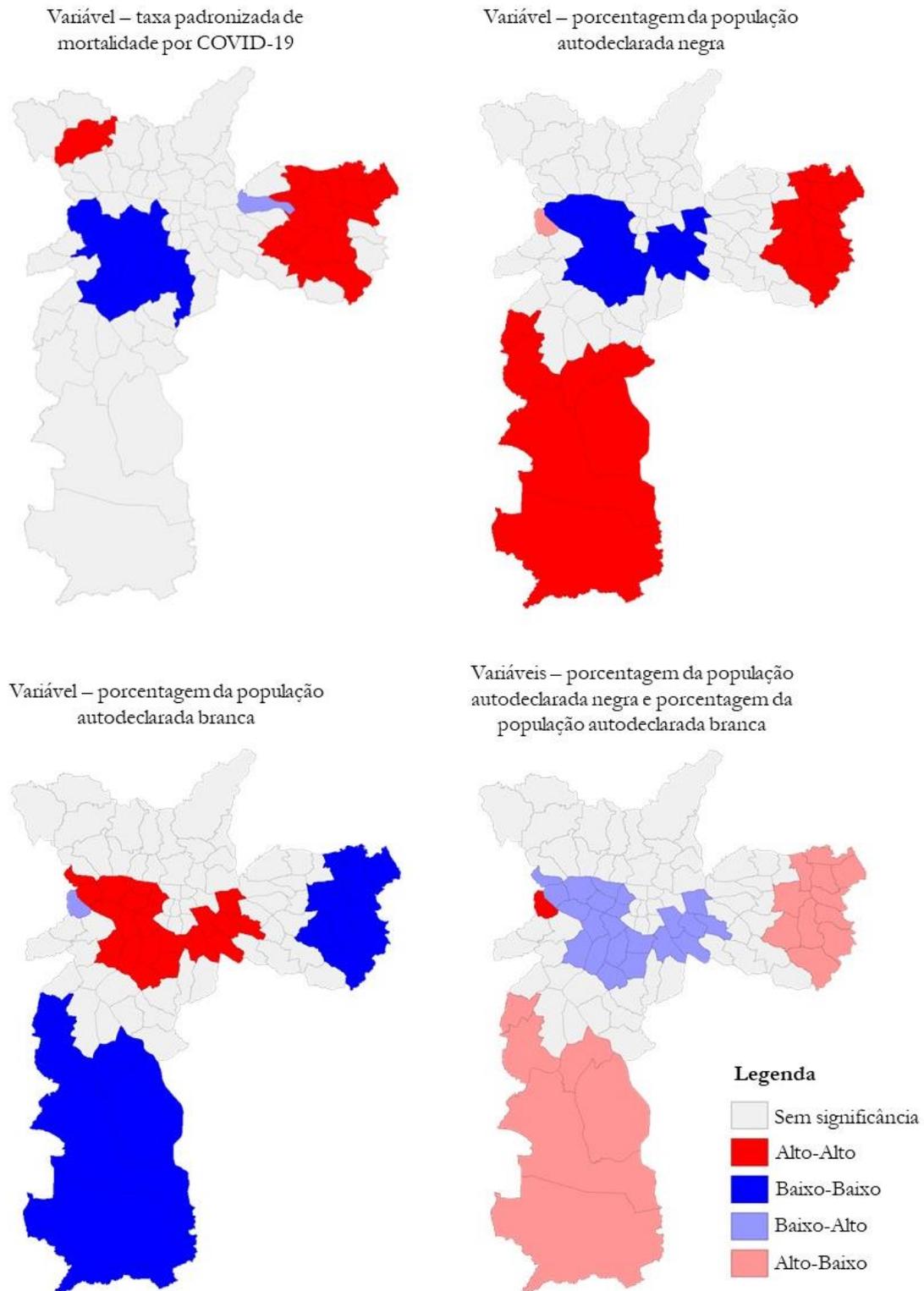
Tabela 04. Resultados dos testes de dependência espacial.

Parâmetros analisados	Taxa padronizada de mortalidade por COVID-19	População autodeclarada negra	População autodeclarada branca	População autodeclarada negra e branca
I de Moran	0,690	0,690	0,666	-0,673
Pseudo p-value	0,001	0,001	0,001	0,001
Multiplicadores de Lagrange (lag)		0.00000	0.00000	0.00000
Teste robusto LM (lag)		0.41780	0.25257	0.32915
Multiplicadores de Langrange (error)		0.00000	0.00000	0.00000
Teste robusto LM (erro)		0.00000	0.00000	0.00000

Fonte: Os autores (2022).

Os valores encontrados para o I de Moran apontam para a autocorrelação presente nas variáveis, sendo todas positivas, à exceção do valor encontrado para o teste bivariado do I de Moran, onde foi encontrada a autocorrelação considerando as duas variáveis independentes juntas. Para este caso, a autocorrelação existe, no entanto, é negativa. Ou seja, onde há maiores porcentagens de população autodeclarada negra há menores porcentagens da população autodeclarada branca e vice-versa. O que pode ser observado na figura 02 onde estão representados os *clusters* formados a partir dos testes univariado e bivariado local de Moran.

Figura 02. Clusters do I de Moran local para as variáveis estudadas.



Fonte: Os autores (2022).

Na figura 02, nota-se a concentração de distritos com altas taxas de mortalidade para COVID-19 na região da zona leste do município. Em contraposição, no centro, zona oeste e parte localizada da zona sul (centro expandido), nota-se um agrupamento dos distritos com menores taxas de mortalidade por COVID-19.

Em relação ao agrupamento da autocorrelação dos distritos de acordo com as variáveis cor autodeclarada pela população, é possível visualizar o óbvio espelhamento das figuras. Os distritos com maiores porcentagens da população autodeclarada negra estão localizados nas regiões sul, sobretudo extremo sul, e extremo leste do município. Em contrapartida as menores porcentagens dessa população são observadas na zona central, oeste e parte da zona sul (centro expandido).

Comparando os resultados da autocorrelação espacial para os índices taxa de mortalidade por COVID-19 e porcentagem da população autodeclarada branca e negra, nota-se que as localidades que apresentam menores taxas de mortalidade por COVID-19 (relação baixo-baixo) possuem maior proporção da população autodeclarada branca (relação alto-alto), enquanto as localidades, principalmente na zona leste, que apresentam maiores taxas de mortalidade por COVID-19 (relação alto-alto), possuem maior proporção de população autodeclarada negra (relação alto-alto).

A partir da observação das figuras é possível identificar que o distrito de Jaguaré se destaca em relação aos seus vizinhos. É o distrito que se destaca como *outlier* na figura 02 (aparece com a relação alto-baixo na imagem que representa a taxa de mortalidade, baixo-alto para variável cor autodeclarada negra e alto-alto quando são consideradas as duas variáveis relacionadas à cor da população). Apesar de estar localizado na zona oeste do município, o distrito é separado de boa parte dos distritos da zona oeste (seus vizinhos imediatos – Alto de Pinheiros e Vila Leopoldina) pelo rio Pinheiros, e está geograficamente mais próximo do município de Osasco (SP). Uma série de estudos tratam do processo de transformação do distrito e áreas adjacentes, sobretudo a partir das operações voltadas à urbanização e reurbanização de favelas na localidade (ver BARIONI et al., 2022 e NAZARETH, 2021). O processo de sobreposição sob uma perspectiva micro escalar, dos novos padrões espaciais aos já consolidados em períodos anteriores (macro padrão centro-periferia), pode ser identificado a partir deste exemplo. Ademais, o exemplo do Jaguaré corrobora com as conclusões de Feitosa et al. (2021) e Marques (2016) sobre o processo de fragmentação e heterogeneização das periferias da metrópole paulista.

Os resultados dos testes de dependência espacial, além de indicarem a importância da incorporação do espaço para ajuste do modelo, também revelaram o *Spatial Error* como o método indicado para realização da regressão espacial. Apesar de os valores de probabilidade dos multiplicadores de Lagrange serem todos inferiores a 0,005, os valores para os testes robustos indicam significância estatística apenas para o *Spatial Error*.

Assim, o *Spatial Error* foi utilizado para realização da regressão espacial e testado para o ajuste do modelo. Da mesma maneira que para a regressão clássica, foram realizados testes individuais, para as duas variáveis independentes, e um teste considerando as duas variáveis como explicativas num mesmo modelo. Os resultados são apresentados na tabela 05.

Tabela 05. Resultados dos testes de regressão espacial *Spatial Error* para as variáveis independentes e a variável dependente: taxa padronizada de mortalidade por COVID-19.

Parâmetros analisados	População autodeclarada negra	População autodeclarada branca	População autodeclarada negra e branca
R2	0,784667	0,767353	0,78760
AIC	981.464	989.196	982.032
p-value População autodeclarada negra	0,000		0.00192
p-value População autodeclarada branca		0,000	0.22956

Fonte: Os autores (2022).

Os resultados do modelo de regressão espacial para as duas variáveis independentes estudadas apresentaram valores melhores em relação aos parâmetros R² e AIC que o modelo de regressão clássico, indicando que o modelo espacial, *Spatial Error*, possui melhor ajuste. A população autodeclarada negra explicou cerca de 78,5% a variância da taxa padronizada de mortalidade por COVID-19, enquanto para a variável população autodeclarada branca esse valor chega a 76,7%, ambos indicam correlação positiva forte entre as variáveis independentes e dependente.

Ao utilizar o modelo de regressão linear múltipla, considerando as duas variáveis como explicativas da taxa padronizada de mortalidade por Covid-19, o valor do R² sobe para 78,8% e o valor de AIC se mantém próximo aos encontrados para as análises de regressão das variáveis independentes de forma individualizada. Ressalta-se que para a análise que leva em consideração as duas variáveis independentes, não foi encontrada significância estatística em relação à população autodeclarada branca (*p-value* equivalente a 0,23). Já para a população autodeclarada negra o *p-value* indica significância estatística (0,002). Esses resultados indicam

que na utilização do modelo *Spatial Error* para cálculo da regressão espacial com duas variáveis independentes e preditoras, os resultados obtidos apenas têm significância estatística para a população autodeclarada negra.

Por fim, o modelo de ajuste local GWR foi aplicado para testes de ajuste local. Os resultados são apresentados na sequência na tabela 06.

Tabela 06. Resultados dos testes de regressão espacial GWR para as variáveis independentes estudadas e a variável dependente taxa padronizada de mortalidade por COVID-19.

Parâmetros analisados	População autodeclarada negra	População autodeclarada branca	População autodeclarada negra e branca
R2	0,8177	0,7948	0,8439
R2 ajustado	0,7704	0,7420	0,7904
AIC	983,23	994,43	980,47

Fonte: Os autores (2022).

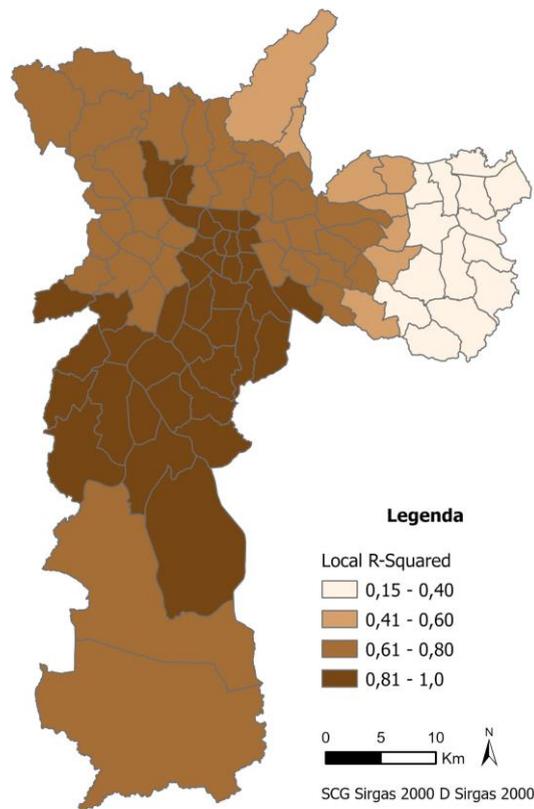
Como pode ser observado, os valores obtidos a partir do método GWR foram os melhores em termos de ajuste dos modelos, seja para as variáveis independentes consideradas de maneira individualizada ou para o modelo que considera as duas porcentagens populacionais como explicativas da taxa de mortalidade por COVID-19.

Considerando apenas a população autodeclarada negra como variável preditora, a variável dependente é explicada em cerca de 77% pela variável independente, no caso do modelo associado à variável população autodeclarada branca esse valor cai para 74%.

Assim como para os resultados dos testes estatísticos apresentados anteriormente, o modelo que apresentou melhores valores de ajuste é o que leva em consideração ambas as variáveis independentes. Ao considerar as porcentagens de populações autodeclaradas negra e branca de maneira combinada, o modelo obteve os maiores valores de R^2 e R^2 ajustado dentre todos os testes estatísticos realizados. Ainda, foi o menor valor de AIC encontrado, ou seja, o “mais simples”.

Na figura 03 são apresentados os locais de maior e menor ajuste do modelo GWR de acordo com o R^2 local.

Figura 03. R² local por distrito do modelo de regressão GWR.



Fonte: Os autores (2022).

Na figura 06 é possível observar onde o modelo foi melhor ajustado, indicando locais onde as variáveis independentes responderam por mais de 80% da variância da taxa padronizada de mortalidade por COVID-19. Para boa parte dos distritos, o modelo teve um bom ajuste e indica uma correlação positiva forte entre as variáveis independentes e dependente estudadas¹. Tem destaque a região central e partes das zonas sul e oeste, para os quais os valores de R² superaram 0,80. Ainda, apresentaram correlação positiva forte entre as variáveis, com valores de R² entre 0,61 e 0,80, distritos localizados na zona norte, oeste e no extremo sul do município, como Parelheiros e Marsilac. Ainda, a correlação foi identificada como forte para alguns dos distritos da zona leste, notadamente os que estão mais próximos à região central.

¹ Foram criadas quatro classes para representação do R². Valores até 0,4 foram considerados como indicativos de uma correlação fraca entre as variáveis; valores entre 0,41 e 0,60 como indicativos de uma correlação moderada; a partir de 0,61 os valores já podem ser interpretados como indicativos de uma correlação forte. A classe de valores maiores que 0,81 foi criada para melhorar a visualização na figura. As classes de visualização foram baseadas no trabalho de Figueiredo Filho e Silva Junior (2009).

Para um grupo de oito distritos, os valores de R^2 encontrados indicaram uma correlação moderada, são dois na zona norte – Jaçanã e Tremembé, este último com parte de seu território abrangido pelo Parque Estadual da Serra da Cantareira. Outros seis estão localizados na zona leste.

Já para quatorze dos distritos localizados na zona leste os valores de R^2 indicaram uma correlação fraca, mas não insignificante, entre as variáveis estudadas. O que aponta para a insuficiência do modelo em prever a variação da taxa padronizada de mortalidade por COVID-19 nessas localidades apenas a partir da composição da população por raça/cor. Esses resultados trazem a reflexão de que as altas taxas de mortalidade padronizada por COVID-19, estão relacionadas a situações habitacionais precárias e a territórios caracterizados pela maior vulnerabilidade social, assim como discutido pela literatura internacional (ver KARAYE; HORNEY, 2020, MAH; ANDREW, 2022; MACKAY et al., 2021, WANG et al., 2022). Assim como concluído por Kim e Bostwick (2020), os resultados apontam que os efeitos desiguais do COVID-19 sobre a população negra refletem as desigualdades sociais e territoriais existentes antes da pandemia.

No entanto, é importante mencionar que os resultados devem ser interpretados com cautela, uma vez que existem limitações próprias da metodologia, como a possível ocorrência de um viés na agregação ou falácia ecológica. Resumidamente, a falácia ecológica refere-se ao fato de que a existência de relações entre variáveis em um nível de agregação pode não se manter em um nível de desagregação ou individual (JARGOWSKY, 2005). Feita as devidas considerações metodológicas, acrescenta-se que o estudo apresenta validade pois os dados são representativos tanto nos aspectos geográficos como demográficos.

Conclusões

Os impactos desiguais da pandemia nos diversos grupos populacionais têm sido documentados pela literatura, destacando que a população negra sofre com as maiores taxas de mortalidade por COVID-19.

Este estudo conclui que existe uma relação positiva e significativa entre as maiores porcentagens de população autodeclarada negra e maiores valores para a taxa padronizada de mortalidade por COVID-19 na cidade de São Paulo (R^2 de 0,52). Os modelos de estatística espacial global e local demonstram um melhor ajuste do modelo de regressão, representado pelo coeficiente de determinação R^2 , respectivamente 0,78 e 0,84, para as variáveis de

composição por raça/cor e as taxas de mortalidade padronizada por COVID-19. O que indica que o espaço é um atributo determinante para relação entre as variáveis.

Os resultados nos permitem refletir que as maiores taxas de mortalidade por COVID-19 estão relacionadas a situações habitacionais precárias, a territórios caracterizados pela maior vulnerabilidade social e à segregação urbana. Em São Paulo, estes territórios também são caracterizados, também, pela maior presença da população autodeclarada negra.

Conhecendo os determinantes sociais e espaciais, políticas públicas podem ser elaboradas tendo como perspectiva não só minimizar os efeitos desiguais na mortalidade por COVID-19 de forma paliativa, alocando maiores recursos humanos e financeiros nestas áreas, mas também, considerando as causas destes efeitos desiguais e buscando políticas públicas mais estruturais que visem reduzir as desigualdades sociais e territoriais.

Referências

ALLEN, Michael Patrick. **Understanding Regression Analysis**. New York and London: Plenum Press, 1997.

ANSELIN, Luc. Lagrange Multiplier Test Diagnostics for Spatial Dependence and Spatial Heterogeneity. **Geographical Analysis**, v. 20, n. 1, p. 1-17, 1988.

ANSELIN, Luc. Local indicators of spatial association-LISA. **Geographical analysis**, v. 27, n. 2, p. 93-115, 1995.

ANDRADE, Lucas Almeida; GOMES, Dharlinton. Soares.; GÓES, Marco Aurélio de Oliveira.; SOUZA, Mércia Simoni Feitosa.; TEIXEIRA, Daniela Cabral Pizzi., RIBEIRO, Caíque Jordan Nunes.; ALVES, José Antônio Barreto; B.; ARAÚJO, Karina Conceição Gomes Machado; SANTOS, Allan Dantas. Surveillance of the first cases of COVID-19 in Sergipe using a prospective spatiotemporal analysis: the spatial dispersion and its public health implications. **Revista da Sociedade Brasileira de Medicina Tropical**, v. 53, 2020.

BAQUI, Pedro; BICA, Ioana; MARRA, Valério; ERCOLE, Ari; VAN DEER SCHAR, Mihaela. Ethnic and regional variations in hospital mortality from COVID-19 in Brazil: a cross-sectional observational study. **The Lancet Global Health**, v. 8, n. 8, p. e1018-e1026, 2020.

BARIONI, Bárbara Cavalcante de Andrade; CARRARO JUNIOR, Hamilton; CAVALCANTE, Danilo de Lima. Análise da exploração imobiliária de autoconstruções em áreas degradadas: Favela Nova Jaguaré. **urbe. Revista Brasileira de Gestão Urbana**, v. 14, p. e20210084, 2021.

BONDUKI, Nabil; ROLNIK, Raquel. Periferia da Grande São Paulo: Reprodução do espaço como expediente de reprodução da força de trabalho. In: MARICATO, E. (Org.). **A produção Capital da casa e da Cidade no Brasil Industrial**. São Paulo: AlfaOmega, 1982.

BRUNSDON, Chris; FOTHERINGHAM, A. Stewart; CHARLTON, Martin E. Geographically weighted regression: a method for exploring spatial nonstationarity. **Geographical analysis**, v. 28, n. 4, p. 281-298, 1996.

CALDEIRA, Teresa Pires do Rio. **Cidade de Muros: crime, segregação e cidadania em São Paulo**. São Paulo: Edusp, 2000

CAMARGO, Cândido Procópio Ferreira. **São Paulo, 1975: crescimento e pobreza**. São Paulo: Loyola, 1976.

CESTARI, Virna Ribeiro Feitosa; FLORENCIO, Raquel Sampaio; SOUSA, George Jó Bezerra; GARCES, Thiago Santos.; MARANHÃO, Thatiana Araújo.; CASTRO, Révia Ribeiro; CORDEIRO, Luana Ibiapina; DAMASCENO, Lara Lúdia Ventura; PESSOA, Vera Lúcia Mendes de Paula; PEREIRA, Maria Lúcia Duarte; MOREIRA, Thereza Maria Magalhães. Vulnerabilidade social e incidência de COVID-19 em uma metrópole brasileira. **Ciência & Saúde Coletiva**, v. 26, p. 1023-1033, 2021.

COY, Martin; PÖHLER, Martin. Gated Communities in Latin American Megacities: Case Studies in Brazil and Argentina. **Environment and Planning B: Urban Analytics and City Science**, v. 29, n. 3, p. 355-370, 2002.

FEITOSA, Flávia da Fonseca; BARROS, Joana; MARQUES, Eduardo César Leão; GIANNOTTI, Mariana. Measuring Changes in Residential Segregation in São Paulo in the 2000s. In: van Ham, M.; Tammaru, T.; Ubarevičienė, R. and Janssen, H. (Org.) **Urban Socio-Economic Segregation and Income Inequality**, Springer, 2021.

FIGUEIREDO FILHO, Dalson Britto; SILVA JUNIOR, José Alexandre. Desvendando os mistérios do coeficiente de correlação de Pearson (r)*. **Revista Política Hoje**, Vol. 18, n. 1, p. 115-146 2009.

FRANÇA, Danilo. Inequalities and Residential Segregation by Race and Class. In: MARQUES, E. C. L. (org.) **São Paulo twenty-first Century spaces, Heterog. Inequalities**. New York: Routledge, 2016, p. 160-174.

IBGE. **Censo demográfico de 2010**. Rio de Janeiro: IBGE, 2010.

IBGE. Desigualdades sociais por cor ou raça no Brasil. **Estudos e Pesquisas - Informação Demográfica e Socioeconômica**, n.41, Rio de Janeiro: IBGE, 2019.

JARGOWSKY, Paul A. The ecological fallacy. **Encyclopedia of social measurement**, v. 1, p. 715-722, 2005.

KARAYE, Ibraheem M.; HORNEY, Jennifer A. The impact of social vulnerability on COVID-19 in the US: an analysis of spatially varying relationships. **American journal of preventive medicine**, v. 59, n. 3, p. 317-325, 2020.

KIM, Sage J.; BOSTWICK, Wendy. Social vulnerability and racial inequality in COVID-19 deaths in Chicago. **Health education & behavior**, v. 47, n. 4, p. 509-513, 2020.

KOWARICK, Lúcio. **A espoliação urbana**. Rio de Janeiro: Paz e Terra, 1979.

KOWARICK, Lúcio. Viver em risco: sobre a vulnerabilidade no Brasil urbano. **Novos Estudos Cebrap**, n. 63, p. 9-30, 2002.

MACKAY, Katharine; AYERS, Chelsea K.; KONDO, Karli K., SAHA, Somnath; ADVANI, Shailesh M.; YOUNG, Sarah; SPENCER, Hunter; RUSEK, Max; ANDERSON, Johanna; VEAZIE, Stephanie; SMITH, Mila; KANSAGARA, Devan. Racial and ethnic disparities in COVID-19–related infections, hospitalizations, and deaths: a systematic review. **Annals of internal medicine**, v. 174, n. 3, p. 362-373, 2021.

MAH, Jasmine C.; ANDREW, Melissa K. Social vulnerability indices: A pragmatic tool for COVID-19 policy and beyond. **The Lancet Regional Health–Europe**, v. 14, 2022.36, p. 163-195, 2022.

MACIEL, Jacques Antonio Cavalcante; CASTRO-SILVA, Igor Iuco; FARIAS, Mariana Ramalho de. Análise inicial da correlação espacial entre a incidência de COVID-19 e o desenvolvimento humano nos municípios do estado do Ceará no Brasil. **Revista brasileira de epidemiologia**, v. 23, 2020.

MARICATO, Ermínia. **A produção Capital da casa e da Cidade no Brasil Industrial**. São Paulo: AlfaOmega, 1979.

MARQUES, Eduardo César Leão. The Social Spaces of the Metropolis in the 2000s. In: MARQUES, E.C.L. (Org.). **São Paulo twenty-first Century spaces, Heterog. Inequalities**. New York: Routledge, 2016.

MAROKO, Andrew R.; NASH, Denis; PAVILONIS, Brian T. COVID-19 and inequity: a comparative spatial analysis of New York City and Chicago hot spots. **Journal of Urban Health**, v. 97, n. 4, p. 461-470, 2020.

NAZARETH, Miguel Bustamante Fernandes. Da favela urbanizada ao urbano no Brasil. **Brazilian Journal of Development**, Curitiba, v.7, n.12, p. 112246-112263, 2021.

RAYMUNDO, Carlos Eduardo; OLIVEIRA, Marcela Cini; ELEUTERIO, Tatiana de Araújo; ANDRÉ, Susana Rosa; SILVA, Marcele Gonçalves; QUEIROZ, Eny Regina da Silva; MEDRONHO, Roberto de Andrade. Spatial analysis of COVID-19 incidence and the sociodemographic context in Brazil. **PLoS One**, v. 16, n. 3, p. e0247794, 2021.

RAFAEL, Ricardo de Mattos Russo, NETO, Mercedes; DEPRET, Davi Gomes; GIL, Adriana Costa; FONSECA, Mary Hellem Silva; SOUZA-SANTOS, Reinaldo. Effect of income on the cumulative incidence of COVID-19: an ecological study. **Revista Latino-Americana de Enfermagem**, v. 28, 2020.

SANTOS, Milton. **Metrópole corporativa fragmentada: o caso de São Paulo**. São Paulo: EDUSP, 1978.

SANTOS, Mércia Pereira Alves dos; NERY, Joilda Silva; GOES, Emanuely Freitas; SILVA, Alexandre da; SANTOS, Andreia Beatriz Silva; BATISTA, Luís Eduardo; ARAÚJO, Edna Maria. População negra e Covid-19: reflexões sobre racismo e saúde. **Estudos Avançados**, v. 34, p. 225-244, 2020.

SÃO PAULO. SECRETARIA MUNICIPAL DA SAÚDE (SMS). Painel Covid-19. SMS, Disponível em: <https://www.prefeitura.sp.gov.br/cidade/secretarias/saude/vigilancia_em_saude/doencas_e_agrivos/coronavirus/index.php?p=31077>. Acesso em 23/05/2022.

SEADE. In: **SEADE Municípios**. Disponível em: <https://municipios.seade.gov.br/>. Acesso em 17/05/2022.

SILVA, L. I. C.; MORAES, E. S.; SANTOS, M. S. COVID-19 e a população negra: desigualdades acirradas no contexto da pandemia. **Revista Thema**, v. 18, p. 301-318, 2020.

TELLES, Eduard E. Residential Segregation by Skin Color in Brazil. **American sociological review**, v. 57, n. 2, p. 186-197, 1992.

WANG, Chuyuang; LI, Ziqi; MATHEUS, Mason Clay; PRAHARAJ, Sarbeswar; KARNA, Brajesh; SOLÍS, Patricia. The spatial association of social vulnerability with COVID-19

Composição populacional por raça/cor e disparidade em taxas de mortalidade por Covid-19 em São Paulo: uma análise estatística e espacial. Agnes Silva de Araujo; Carina Sernaglia Gomes.

prevalence in the contiguous United States. **International Journal of Environmental Health Research**, v. 32, n. 5, p. 1147-1154, 2022.

YANCY, Clyde W. COVID-19 and African Americans. **JAMA: The Journal of the American Medical Association**, v. 323, n. 19, p. 1891-1892, 2020.

Recebido em: julho de 2023
Aceito em: setembro de 2024